VI. - Erreurs - Estimations Intervalles de sécurité

Erreurs maxima

Nous sommes assez souvent renseignés sur l'erreur maxima à redouter dans une observation. Nous notons par exemple que le diamètre d'un arbre est de 75 cm : en réalité, nous savons qu'il est compris entre 72,5 et 77,5 cm. On traduira l'erreur maxima à craindre en écrivant que le diamètre en cause est de 75 ± 2,5 cm. C'est une question de précision de mesure.

La même notation peut être employée pour définir la variabilité d'une grandeur. La hauteur moyenne d'une plantation est de 9 m, mais certains sujets atteignent jusqu'à 11 m alors que les plus petits ont seulement 7 m. Cette variabilité se traduit en écrivant que la hauteur est de 9 ± 2 m.

Supposons qu'à une longueur caractérisée par l'expression : 34,1 ± 0,05 cm, nous ayons à ajouter une autre longueur qui sera par exemple 27,3 ± 0,05 cm. Le total s'écrira : 61,4 ± 0,10 cm et on entendra par là qu'il est compris entre 61,3 et 61,5 cm.

On a vu par cet exemple que si ε est l'erreur pouvant affecter chacune des deux mesures à additionner, la somme comportera elle-même une erreur maxima $2\,\varepsilon$ et d'une façon générale la somme de N mesures pouvant comporter une erreur ε pourra comporter elle-même une erreur $N\,\varepsilon$.

Quand on soustrait l'une de l'autre deux mesures, la différence sera elle-même affectée de la même erreur maxima que le total, soit $2\,\varepsilon$. Si la production d'un peuplement est connue par la différence de 2 volumes comportant chacun une erreur de 10 m^3, le résultat risque de comporter une erreur de 20 m^3.

Ayant effectué un certain nombre de mesures, nous sommes amenés à calculer la moyenne des mesures faites. La moyenne peut être affectée d'une erreur ε égale à celle que comporte chacune des mesures. Une moyenne possède donc le même degré de précision que chacune des mesures ayant servi à l'établir.

Enfin, lorsqu'il s'agit d'un produit, on envisage les erreurs relatives. Si les deux facteurs A et B d'un produit comportent chacun une erreur de 5 %, le produit AB sera compris entre AB — 10 % AB et AB + 10 % AB. D'une façon générale, si λ est l'erreur relative sur l'un des facteurs, le produit pourra s'écrire : $AB\,(1 \pm 2\,\lambda)$.

Dans tous les cas que nous venons d'envisager, les écarts susceptibles de se produire sont indiqués sans ambiguïté par le signe ±.

Cependant, il est essentiel de bien spécifier qu'on a entendu expliciter ainsi, soit la précision de la mesure s'il s'agit d'une matière inerte, soit la variabilité biologique d'une grandeur étudiée, mais de toute façon un écart maximum.

Ecart-type

En réalité, l'erreur-maxima n'est pas toujours la plus intéressante à considérer. Les écarts de grande amplitude n'ont pas grande importance s'ils sont exceptionnels. C'est donc plutôt sur la dispersion autour de la moyenne que nous désirons être renseignés.

Ainsi qu'on l'a vu dans le cas des courbes en cloche, l'écart-type est tel qu'il définit 2 limites entre lesquelles sont situées 68 % des observations (chapitre III, fig. 8).

Le double de l'écart-type définit des limites plus amples : entre elles, sont situées 95 % des observations.

Reprenons l'exemple du peuplement qui a fait l'objet du tableau n° 1 (chapitre IV, p. 49). Les calculs auxquels nous avons procédé nous ont permis de réduire les observations aux deux paramètres suivants :

Circonférence moyenne	7,21 dm
Ecart-type	2,23

A la condition de prévenir que le signe ± précède la valeur de l'écart-type, on pourra résumer les données dans la formule :

$$7{,}21 \pm 2{,}23$$

Le lecteur prévenu peut alors interpréter que : 95 % des arbres ont une circonférence entre 2,75 et 11,67,
et que 68 % sont entre 4,98 et 9,44.

Un simple examen du tableau n° 1 nous montre que ces prévisions sont à peu près réalisées. Il n'y a que 16 tiges extérieures au premier intervalle qui englobe donc

$$\frac{607 - 16}{607}$$

soit 97 % des tiges.

A l'extérieur du 2e intervalle, on peut considérer qu'il y a les classes 3, 4 et la moitié de la classe 5, soit :

$$24 + 55 + 35 = 114$$

ainsi que les classes 10 et au-dessus, soit

$$51 + 26 + 9 + 5 + 2 = 93$$

Au total, 207 sur 607, soit 34 %, ce qui en met 66 % à l'intérieur du 2e intervalle.

Erreur à craindre correspondant à une probabilité donnée. — La norme X 06-001 (Terminologie des erreurs de mesure) définit comme suit l'erreur à craindre correspondant à une probabilité donnée :

Si la loi de probabilité est celle de Laplace (loi normale), les erreurs moyennes quadratiques — correspondant à l'écart-type — positive et négative sont les erreurs à craindre correspondant à la probabilité 0,6826...

La recommandation suivante de la norme est particulièrement importante :

« Quand un résultat est donné sous la forme $A \pm \varepsilon$, la signification de la quantité ε doit toujours être précisée (par exemple erreur à craindre correspondant à une probabilité donnée).

L. SCHAEFFER.